Robust Distributed Averaging on Networks with Adversarial Intervention

Ali Khanafer, Behrouz Touri, and Tamer Başar

Abstract— **We study the interaction between a network designer and an adversary over a dynamical network. The network consists of nodes performing continuous-time distributed averaging. The goal of the network designer is to assist the nodes reach consensus by changing the weights of a limited number of links in the network. Meanwhile, an adversary strategically disconnects a set of links to prevent the nodes from converging. We formulate two problems to describe this competition where the order in which the players act is reversed in the two problems. We utilize Pontryagin's Maximum Principle (MP) to tackle both problems and derive the optimal strategies. Although the canonical equations provided by the MP are intractable, we provide an alternative characterization for the optimal strategies that highlights a connection with potential theory. Finally, we provide a sufficient condition for the existence of a saddle-point equilibrium (SPE) for this zero-sum game.**

I. INTRODUCTION

Various physical and biological phenomena where global patterns of behaviour stem from local interactions have been modelled using linear averaging dynamics. In such dynamics an agent updates its value as a linear combination of the values of its neighbors. Averaging dynamics is the basic building block in many multi-agent systems, and it is widely used whenever an application requires multiple agents, who are graphically constrained, to synchronize their measurements. Examples include formation control, coverage, distributed estimation and optimization, and flocking [1]–[3]. Besides engineering, linear averaging finds applications in other fields. For instance, social scientists use averaging to describe the evolution of opinions in networks [4].

In practice, communication among agents is prone to different non-idealities which can affect the convergence properties of distributed algorithms. Transmission delays [5], noisy links [6], [7], and quantization [8] are some examples of non-idealities that are due to the physical nature of the application. In addition to physical restrictions, researchers have also studied averaging dynamics in the presence of malicious nodes in the network [9], [10]. Various algorithms that guarantee resilience against node failures have been proposed in the literature [11].

In [12], we explored the effect of an external adversary who attempts to prevent the nodes from reaching consensus by disconnecting certain links in the network. We derived his

optimal strategy and demonstrated that it admits a potential-theoretic analogy. Here, we introduce a network designer that attempts to counter the effect of the adversary and help the nodes reach consensus by changing the weights of certain links. Both the adversary and the designer are constrained by their physical properties, e.g., battery life and communication range. To capture such constraints, we allow the adversary and the designer to affect only a fixed number of links.

Such an interaction between a network designer and an adversary can occur in various practical applications. For example, in a wireless network, the adversary can be a jammer who is capable of breaking links by injecting high noise signals that disrupt the communication among nodes. The link weights in such a network represent the capacities of the corresponding links. The designer can modify the capacity of a certain link using various communication techniques such as introducing parallel channels between two nodes as in orthogonal frequency division multiple access (OFDMA) networks [13].

Our model is different from the current literature in two ways: (i) the adversary and the designer compete over a dynamical network. This is different from the problems studied in the computer science and economics communities where the network is usually static [14]; (ii) the players in our model are constrained and cannot operate with an infinite budget. This enables us to model practical scenarios more closely rather than allowing the malicious behaviour to be unrestricted as in [9], [15], [16].

The main goal of this work is to derive optimal strategies for the designer and the adversary who have conflicting objectives. Because the order in which the players act affects the resulting utilities, we formulate two problems based on the order of play, allowing each player to have the *first-move-advantage* in one of the problems. When the adversary is allowed to play first, he is capable of restricting the available actions of the designer since some links will disappear from the network. Hence, if we were to cast a zero-sum game between the players, one should not expect the existence of a saddle-point equilibrium (SPE) in pure strategies. The question we would like to answer is then: *are there scenarios where the order of play does not affect the utilities of the players, which leads to the existence of an SPE?*

Accordingly, the contributions of this paper are as follows.

- We capture the interaction between the designer and the adversary by formulating two problems. In the min–max problem, the designer declares a strategy first to which the adversary reacts by its optimal response. The second problem is a max–min one where the order of play is reversed. We derive the optimal strategies in both

problems by employing the Maximum Principle (MP).

- We provide a method to compute the optimal strategies without requiring the adjoint equations to be solved. This method provides a new characterization for the optimal strategies in terms of potential-theoretic quantities.
- We derive a sufficient condition guaranteeing the existence of an SPE.

The rest of the paper is organized as follows. In Section II, we formulate and provide the preliminaries of the min–max and max–min problems. We analyze both problems in Section III and derive their solutions. An implementation method is provided in Section IV. We provide a sufficient condition for the existence of an SPE in Section V, and we conclude the paper in Section VI. An Appendix contains two lemmas used in the derivation of the sufficient condition.

Notation: We denote the set of edges in a graph $\mathcal{G}$ by $\mathcal{E}(\mathcal{G})$. When clear from the context, we will drop the argument of any set defined on a graph. To emphasize the effect of link removal by the adversary, we will sometimes write $\mathcal{G}(u(t))$ to denote the graph resulting after the adversary acts at time t. We will use $\sum_{j>i}(.)$ to mean $\sum_{j=2}^{n} \sum_{i=1}^{j-1}(.)$.

II. PROBLEM DESCRIPTION

Consider a connected network of n nodes described by an undirected graph $\mathcal{G}$ with n vertices and m edges. We will denote an edge in $\mathcal{E}(\mathcal{G})$ between nodes i and j by $\{i, j\}$. The value, or state, of the nodes at time instant t is given by $x(t) = [x_1(t), ..., x_n(t)]^T$. The nodes start with an initial value $x(0) = x_0$, and they are interested in computing the average of their initial measurements $x_{avg} = \frac{1}{n}\sum_{i=1}^{n} x_i(0)$ via local averaging. We consider the continuous-time averaging dynamics given by

$$\dot{x}(t) = A(t)x(t), \quad x(0) = x_0, \qquad (1)$$

where the matrix $A(t)$ has the following properties

$$A(t) = A(t)^T, \quad A_{ij}(t) \geq 0 \quad i \neq j, \quad A(t)\mathbf{1} = 0.$$

Define $\bar{x} = \mathbf{1}x_{avg}$ and let $M = \frac{\mathbf{1}\mathbf{1}^T}{n}$. A well-known result states that, given the above assumptions, the nodes will reach consensus as $t \rightarrow \infty$, i.e., $\lim_{t\to\infty} x(t) = \bar{x}$ [1]. In the absence of the adversary and the network designer, $A(t)$ is time-invariant with $A_{ij}(t) = a_{ij}$. To achieve their respective objectives, the designer and the adversary control the elements of $A(t)$ as we explain next.

The adversary attempts to slow down convergence by breaking at most $\ell \leq m$ links at each time instant. Let $u_{ij}(t) \in \{0, 1\}$ be the weight the adversary assigns to link $\{i, j\}$. He breaks link $\{i, j\}$ at time t when $u_{ij}(t) = 1$. His control is given by $u(t) = [u_{12}(t), u_{13}(t), ..., u_{1n}(t), u_{23}(t), ..., u_{(n-1)n}(t)]^T$. We will denote the number of links the adversary breaks at time t by $N_u(t)$. In accordance with the above, the strategy space of the adversary is

$$\mathcal{U} = \left\{ u \in \{0, 1\}^{\binom{n}{2}} : N_u \leq \ell \right\}.$$

Meanwhile, the network designer attempts to accelerate convergence by controlling the weights of the edges. Let $v(t) = [v_{12}(t), v_{13}(t), ..., v_{1n}(t), v_{23}(t), ..., v_{(n-1)n}(t)]^T$ be the control of the designer. The designer can change the weight of a given link by adding $v_{ij}(t)$ to its weight a_{ij}. We assume that $v_{ij}(t) \in [0, b]$ and that the number of links the designer modifies $N_v(t)$ is at most ℓ. The strategy set of the designer is therefore

$$\mathcal{V} = \left\{ v \in [0, b]^{\binom{n}{2}} : N_v \leq \ell \right\}.$$

If $\{i, j\} \notin \mathcal{E}$, then $u_{ij}(t) = v_{ij}(t) = 0$ for all t. Given the above definitions, we can write the $\{i, j\}$-th element of the matrix $A(u(t), v(t))$ as

$$A_{ij}(u(t), v(t)) = (a_{ij} + v_{ij}(t)) \cdot (1 - u_{ij}(t)).$$

Implicit in this definition is the assumption that the effect of the adversary and the designer on the links lasts only instantly, and the weights of the links are reset to a_{ij} at the next time instant. This is motivated by applications where the state of a channel deteriorates only when the adversary is acting, e.g., wireless communications where the effect of the jammer is not permanent.

Define the following functional:

$$J(u, v) = \int_0^T k(t) |x(t) - \bar{x}|^2 dt,$$

where the kernel $k(t)$ is positive and integrable over $[0, T]$. The utility of the adversary is then given by $J(u, v)$, and that of the designer is $-J(u, v)$. We will study the following two problems. In the first problem, the adversary acts first by selecting the links he is interested in breaking. Then, the network designer optimizes his choices over the resulting graph $\mathcal{G}(u(t))$. More formally, we have

$$\max_{u(t)\in\mathcal{U}} \min_{v(t)\in\mathcal{V}} J(u, v)$$
$$\text{s.t.} \qquad \dot{x}(t) = A(u(t), v(t))x(t), \quad x(0) = x_0.$$

In the second problem, the order is reversed, and the designer acts first:

$$\min_{v(t)\in\mathcal{V}} \max_{u(t)\in\mathcal{U}} J(u, v)$$
$$\text{s.t.} \qquad \dot{x}(t) = A(u(t), v(t))x(t), \quad x(0) = x_0.$$

In a computer network, where consensus could model the spread of a virus, the max–min problem allows the network architect (who is the maximizer here) to implement networks that are robust against a strategic virus diffusion. The min–max problem finds applications in army combat situations where the designer (the minimizer) attempts to counter the attacks of the enemy to disrupt the communications among agents. In both problems, we make the following assumption:

Assumption 1: The initial matrix $A(0, 0)$, the state $x(t)$, and the value b are common information to both players. The following two remarks are in order.

Remark 1: *(Problem Complexity)* Let us consider the problem of the adversary for a given strategy of the designer.

Assume that the adversary can act in K time instances over the interval $[0, T]$. Then, for small ℓ, the total number of links that need to be tested in a brute-force approach is

$$\binom{m}{\ell}^K \approx m^{\ell K}. \quad (2)$$

Clearly, the brute-force approach incurs exponential-time complexity, especially if any of m, ℓ, k becomes large. The same argument applies to the network designer.

Remark 2: *(Non-Rectangular Strategies)* When the strategy sets are rectangular, i.e., the strategy of one player does not restrict the strategy space of the other, the following holds

$$\underline{V} = \max_{u(t) \in \mathcal{U}} \min_{v(t) \in \mathcal{V}} J(u, v) \le \min_{v(t) \in \mathcal{V}} \max_{u(t) \in \mathcal{U}} J(u, v) = \overline{V}, \quad (3)$$

where $\underline{V}, \overline{V}$ are called the lower and upper values of the game, respectively. An SPE in pure strategies exists when $\underline{V} = \overline{V}$. When the strategy spaces are non-rectangular, one should *not* expect the existence of an SPE in pure-strategies nor the order in (3). In the max–min problem, the adversary restricts the strategy space of the designer since $u(t)$ removes links from $\mathcal{G}$. Hence, the max–min and min–max problems should be studied separately.

III. Adversary vs. Network Designer: Two Problems

We will now present the solutions for the two problems presented above. To arrive at the optimal strategies of the players, we employ the MP. In Section IV, we prove that $N_u^\star = N_v^\star = \ell$; we will use this fact in the proofs of this section. In what follows, we will often drop the time index and other arguments for notational simplicity.

A. The Min–Max Problem

The Hamiltonian associated with the max–min problem is:

$$H(x, p, u, v) = k(t) |x(t) - \bar{x}|^2 + p^T(t) A(t) x(t).$$

The first-order necessary conditions for optimality are:

$$
\begin{aligned}
\dot{p} &= -\frac{\partial}{\partial x} H \\
&= -2k(x - \bar{x}) - Ap, \quad p(T) = 0 \quad (4) \\
\dot{x} &= Ax, \quad x(0) = x_0 \quad (5) \\
(u^\star, v^\star) &= \arg\max \arg\min \{H : u \in \mathcal{U}, v \in \mathcal{V}\},
\end{aligned}
$$

where $x, p \in C^1[0, T]$, the space of continuously differentiable functions over $[0, T]$.

To find the optimal strategies, let us first write

$$
\begin{aligned}
p^T A x &= \sum_{i=1}^n p_i \left(\sum_{j=1}^n A_{ij} x_j \right) \\
&= \sum_{i=1}^n p_i \left(-\sum_{j=1, j\ne i}^n A_{ij} x_i + \sum_{j=1, j\ne i}^n A_{ij} x_j \right) \\
&= \sum_{j>i} (a_{ij} + v_{ij})(1 - u_{ij})(p_j - p_i)(x_i - x_j).
\end{aligned}
$$

Define the function

$$f_{ij} = (p_j - p_i)(x_i - x_j),$$

and let us write

$$
\begin{aligned}
\max_{u \in \mathcal{U}} \min_{v \in \mathcal{V}} H &= \max_{u \in \mathcal{U}} \min_{v \in \mathcal{V}} k |x - \bar{x}|^2 + p^T A x \\
&= k |x - \bar{x}|^2 + \max_{u \in \mathcal{U}} \min_{v \in \mathcal{V}} \sum_{j>i} A_{ij} f_{ij}. \quad (6)
\end{aligned}
$$

Note that we cannot decouple the maximization (or minimization) into $\binom{n}{2}$ maximization problems, each corresponding to a link or a single term inside the double summation. This is due to the constraint on the number of links that can be targeted by the players. Before finding the optimal strategies, let us first define the following sets. Let $\mathcal{F}(\mathcal{G}') = \{f_{ij} < 0 : \{i, j\} \in \mathcal{E}(\mathcal{G}')\}$ for some graph $\mathcal{G}'$. It should be noted that although $\{i, j\}$ and $\{j, i\}$ belong to $\mathcal{E}(\mathcal{G}')$, we only include f_{ij} once in $\mathcal{F}(\mathcal{G}')$, i.e., we treat $\{i, j\}$ and $\{j, i\}$ as one link. This applies to all the definitions to follow. Let $\mathcal{F}_\pi(\mathcal{G}')$ be a ranking of the set $\mathcal{F}(\mathcal{G}')$ such that $\mathcal{F}_{\pi,1}(\mathcal{G}') \le \ldots \le \mathcal{F}_{\pi,|\mathcal{F}(\mathcal{G}')|}(\mathcal{G}')$. Further, let $\mathcal{F}_\pi^\ell(\mathcal{G}')$ be the set containing the smallest ℓ values in $\mathcal{F}_\pi(\mathcal{G}')$. It is understood that if $\mathcal{F}_\pi(\mathcal{G}')$ contains fewer than ℓ negative elements, $\mathcal{F}_\pi^\ell$ will contain all those elements. Define the set operator $\Phi_i : S(\mathcal{G}') \to \mathcal{E}(\mathcal{G}')$ that returns the links in $\mathcal{E}(\mathcal{G}')$ that correspond to the elements of $S(\mathcal{G}')$. Also, define $\Phi_i : S(\mathcal{G}') \to \mathcal{E}(\mathcal{G}')$ that returns the links in $\mathcal{E}(\mathcal{G}')$ corresponding to the *smallest* i elements of the set $S(\mathcal{G}')$. When $|S(\mathcal{G}')| < i$, $\Phi_i(S(\mathcal{G}')) = \Phi(S(\mathcal{G}'))$. We also adopt the convention $\Phi_0(.) = \{\emptyset\}$. We can then write

$$\mathcal{F}_\pi^\ell(\mathcal{G}') = \Phi_\ell \left(\{f_{ij} \le \mathcal{F}_{\pi, \ell+1}(\mathcal{G}') : \{i, j\} \in \mathcal{E}(\mathcal{G}')\} \right).$$

Define the sets $\mathcal{D}_1(\mathcal{G}') = \{a_{ij} f_{ij} < 0 : \{i, j\} \in \mathcal{E}(\mathcal{G}')\}$, $\mathcal{D}_2(\mathcal{G}') = \{(a_{ij} + b) f_{ij} < 0 : \{i, j\} \in \mathcal{E}(\mathcal{G}')\}$, and $\mathcal{D}(\mathcal{G}') = \mathcal{D}_1(\mathcal{G}') \cup \mathcal{D}_2(\mathcal{G}')$. Also, let $\mathcal{D}_\pi(\mathcal{G}')$ be a ranking of $\mathcal{D}(\mathcal{G}')$ such that $\mathcal{D}_{\pi,1}(\mathcal{G}') \le \ldots \le \mathcal{D}_{\pi,|\mathcal{D}(\mathcal{G}')|}(\mathcal{G}')$. Finally, define

$$
\begin{aligned}
\mathcal{D}_\pi^\ell(\mathcal{G}') = \Phi_\ell(\{&a_{ij} f_{ij} \le \mathcal{D}_{\pi, \ell+1}(\mathcal{G}') \text{ or} \\
&(a_{ij} + b) f_{ij} \le \mathcal{D}_{\pi, \ell+1}(\mathcal{G}') : \{i, j\} \in \mathcal{E}(\mathcal{G}')\}).
\end{aligned}
$$

The following theorem specifies the optimal strategies of the adversary and the designer.

Theorem 1: The optimal strategies for the network designer and the adversary in the max–min problem are

$$
v_{ij}^\star(t) = \begin{cases}
b, & \{i, j\} \in \mathcal{F}_\pi^\ell(\mathcal{G}^\star) \\
0, & \{i, j\} \notin \mathcal{F}_\pi^\ell(\mathcal{G}^\star) \text{ or } f_{ij} > 0 \\
[0, b], & \text{otherwise}
\end{cases}
$$

$$
u_{ij}^\star(t) = \begin{cases}
1, & \{i, j\} \in \mathcal{D}_\pi^\ell(\mathcal{G}) \\
0, & \{i, j\} \notin \mathcal{D}_\pi^\ell(\mathcal{G}) \text{ or } f_{ij} > 0 \\
\{0, 1\}, & \text{otherwise}
\end{cases}
$$

where $\mathcal{G}^\star = \mathcal{G}(u^\star)$ is the graph resulting after the adversary applies $u^\star$.

Proof: For both strategies, when $f_{ij} = 0$, the MP does not tell us what the optimal action is. Hence, both players can select any action from their strategy spaces in this case, provided they do not exceed their budget constraints. Let us

start by deriving $v^\star$. By (6), the terms $a_{ij}f_{ij}$ do not affect the minimization problem. Also, note that $\mathcal{G}^\star$ contains the links for which $u_{ij}^* = 0$ only. Then, we can write

$$\operatorname*{argmin}_{v \in \mathcal{V}} \sum_{j>i} A_{ij}f_{ij} = \operatorname*{argmin}_{v \in \mathcal{V}} \sum_{\substack{j>i \\ \{i,j\} \in \mathcal{E}(\mathcal{G}^\star)}} v_{ij}f_{ij}. \quad (7)$$

Now consider the sum on the right hand side. Being a minimizer, the designer should not amplify the positive f_{ij}'s; hence, he must set $v_{ij}^\star = 0$ when $f_{ij} > 0$. Then, by recalling the definition of $\mathcal{F}_\pi^\ell(\mathcal{G}^\star)$, we obtain

$$\sum_{\substack{j>i \\ \{i,j\} \in \mathcal{E}(\mathcal{G}^\star)}} v_{ij}f_{ij} = \sum_{\substack{j>i \\ \{i,j\} \in \mathcal{F}_\pi^\ell(\mathcal{G}^\star)}} v_{ij}f_{ij} + \sum_{\substack{j>i \\ \{i,j\} \notin \mathcal{F}_\pi^\ell(\mathcal{G}^\star)}} v_{ij}f_{ij}$$

$$\geq \sum_{\substack{j>i \\ \{i,j\} \in \mathcal{F}_\pi^\ell(\mathcal{G}^\star)}} bf_{ij}.$$

Note that the last inequality can indeed be achieved because $|\mathcal{F}_\pi^\ell(\mathcal{G}^\star)| \leq \ell$. Hence, the optimal $v^\star$ is as claimed.

In this problem, the adversary has the first-mover-advantage and needs to dispose of the links that can reduce his utility. Because the adversary attempts to maximize the minimum of $p^T A x$, it follows that $u_{ij}^* = 0$ when $f_{ij} > 0$. Let us now consider the links with $f_{ij} < 0$. The adversary knows that the designer attempts to make the negative f_{ij}'s smaller by adding b to the corresponding edge weights. However, we cannot rule out the possibility that $(a_{lk}+b)f_{lk} > a_{ij}f_{ij}$. Hence, the adversary is not only interested in finding the smallest negative $(a_{ij}+b)f_{ij}$'s, but also needs to consider the $a_{ij}f_{ij}$'s themselves. It follows that the adversary needs to find the terms that can become very small (negative) and set $u_{ij} = 1$ to the corresponding links. But those links are exactly the ones included in $\mathcal{D}_\pi^\ell(\mathcal{G})$. Formally, and using the notation $A_{ij}^* = (a_{ij}+v_{ij}^\star)(1-u_{ij})$, we have

$$\sum_{j>i} A_{ij}^\star f_{ij} = \sum_{\substack{j>i \\ \{i,j\} \in \mathcal{D}_\pi^\ell(\mathcal{G})}} A_{ij}^\star f_{ij} + \sum_{\substack{j>i \\ \{i,j\} \notin \mathcal{D}_\pi^\ell(\mathcal{G})}} A_{ij}^\star f_{ij}$$

$$\leq \sum_{\substack{j>i \\ \{i,j\} \notin \mathcal{D}_\pi^\ell(\mathcal{G})}} (a_{ij}+v_{ij}^\star)f_{ij},$$

where the last inequality follows from the fact that $f_{ij} < 0$ for all $\{i,j\} \in \mathcal{D}_\pi^\ell(\mathcal{G})$. This confirms that $u^\star$ is as claimed. ∎

B. The Min–Max Problem

We now study the min–max problem. Similar to the above, we can write

$$\min_{v \in \mathcal{V}} \max_{u \in \mathcal{U}} H = k\left|x-\bar{x}\right|^2 + \min_{v \in \mathcal{V}} \max_{u \in \mathcal{U}} \sum_{j>i} A_{ij}f_{ij}.$$

We again note that one needs to be careful to observe that the minimization (or maximization) problem does not decouple due to the constraint on the controls of the two players. Although the adversary's strategy can be derived using an approach similar to the previous problem, the designer's strategy requires more care. Let $\mathcal{L}(v) = \{(a_{ij}+v_{ij})f_{ij} <$

0: **input:** a strategy v with $N_v = 0$
1: **for** $i = \ell \downarrow 1$
2:     **if** $\exists \mathcal{S} \subset \mathcal{P}(0), |\mathcal{S}| = i, \mathcal{L}_{\pi,i}^\ell(0) \notin \mathcal{L}_\pi^\ell(v_\mathcal{S}(b))$
3:        Set $v_{ij} = b, \forall \{i,j\} \in \Phi(\mathcal{S}) \cup \Phi_{\ell-i}\left(\overline{\mathcal{P}}(v_\mathcal{S}(b))\right)$.
4:        **Exist** for loop.
5:     **end**
6: **end**
7: **if** $N_v = 0$
8:     Set $v_{ij} = b$ for all $\{i,j\} \in \Phi_\ell\left(\overline{\mathcal{P}}(0)\right)$.
9: **end**

$0 : \{i,j\} \in \mathcal{E}(\mathcal{G})\}$. Define $\mathcal{L}_\pi(v)$ to be a ranking of the set $\mathcal{L}(v)$ such that $\mathcal{L}_{\pi,1}(v) \leq \ldots \leq \mathcal{L}_{\pi,|\mathcal{L}|}(v)$, and the set

$$\mathcal{L}_\pi^\ell(v) = \Phi_\ell\left(\{(a_{ij}+v_{ij})f_{ij} \leq \mathcal{L}_{\pi,\ell+1}(\mathcal{G}) : \{i,j\} \in \mathcal{E}(\mathcal{G})\}\right).$$

As an abuse of notation, we let $\mathcal{L}_{\pi,k}^\ell(v)$ correspond to the *value* $(a_{ij}+v_{ij})f_{ij}$ associated with the link $\{i,j\} \in \mathcal{L}_\pi^\ell(v)$. We assume that $\mathcal{L}_{\pi,1}^\ell(v) \geq \ldots \geq \mathcal{L}_{\pi,\ell}^\ell(v)$. Further, define the sets $\mathcal{P}(v) = \{a_{ij}f_{ij} < 0 : \{i,j\} \notin \mathcal{L}_\pi^\ell(v)\}$ such that $\mathcal{P}_1(v) \leq \ldots \leq \mathcal{P}_{|\mathcal{P}|}(v)$ and $\overline{\mathcal{P}}(v) = \{f_{ij} < 0 : \{i,j\} \notin \mathcal{L}_\pi^\ell(v)\}$ such that $\overline{\mathcal{P}}_1(v) \leq \ldots \leq \overline{\mathcal{P}}_{|\overline{\mathcal{P}}|}(v)$. We also define

$$[v_\mathcal{S}(b)]_{ij} = \begin{cases} b, & \{i,j\} \in \mathcal{S} \\ 0, & \{i,j\} \notin \mathcal{S} \end{cases}$$

The following theorem specifies the optimal strategies of both players.

Theorem 2: The optimal strategy for the adversary in the min–max problem is

$$u_{ij}^\star(t) = \begin{cases} 1, & \{i,j\} \in \mathcal{L}_\pi^\ell(v^\star) \\ 0, & \{i,j\} \notin \mathcal{L}_\pi^\ell(v^\star) \text{ or } f_{ij} > 0 \\ \{0,1\}, & \text{otherwise} \end{cases}$$

For the network designer, the optimal strategy is to run Algorithm I and set $v_{ij}^\star = 0$ when $f_{ij} > 0$ and $v_{ij}^\star \in [0,b]$ if $f_{ij} = 0$.

Proof: The same reasoning as in Thm 1 applies to the case when $f_{ij} \geq 0$. We start by deriving $u^\star$. By recalling the definition of $\mathcal{L}_\pi^\ell(v^\star)$, we can write

$$\sum_{j>i} A_{ij}^\star f_{ij} = \sum_{\substack{j>i \\ \{i,j\} \in \mathcal{L}_\pi^\ell(v^\star)}} A_{ij}^\star f_{ij} + \sum_{\substack{j>i \\ \{i,j\} \notin \mathcal{L}_\pi^\ell(v^\star)}} A_{ij}^\star f_{ij}$$

$$\leq \sum_{\substack{j>i \\ \{i,j\} \notin \mathcal{L}_\pi^\ell(v^\star)}} (a_{ij}+v_{ij}^\star)f_{ij},$$

where the last inequality follows from the fact that $f_{ij} < 0$ for all $\{i,j\} \in \mathcal{L}_\pi^\ell(v^\star)$. This confirms that $u^\star$ is as claimed. Also, note that the last inequality can be achieved because $|\mathcal{L}_\pi^\ell(v^\star)| \leq \ell$ by construction.

As for the designer, he can now exploit the first-mover-advantage since he is acting first. The main goal of the designer in this case is to try to protect the links in $\mathcal{L}_\pi^\ell(0)$. These links will be disconnected by the adversary unless the designer is able to modify the ranking of the links such that
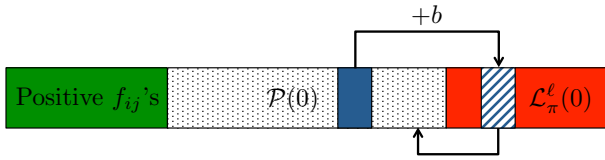
Fig. 1. A demonstration of how the designer can alter the ranking of the f_{ij}'s to improve his utility.

the links (or a subset of them) in $\mathcal{L}_\pi^\ell(0)$ are not in $\mathcal{L}_\pi^\ell(v^\star)$. In essence, this is what Algorithm I attempts to achieve. This is demonstrated in Fig. 1. Being the lowest negative value, and hence the link both the adversary and the designer are interested in, let us explore how the designer can push $\mathcal{L}_{\pi,\ell}^\ell(0)$ higher in the ranking of the link values. The designer can achieve this if under some strategy v, the value $\mathcal{L}_{\pi,\ell}^\ell(0)$ is no longer among the lowest ℓ negative values; in other words, the designer can alter the ranking if there is a set of ℓ links $\mathcal{S} \subset \mathcal{P}(0)$ such that when he sets $v_{ij} = b$ for all links in $\mathcal{S}$, there will be ℓ values that are smaller than $\mathcal{L}_{\pi,\ell}^\ell(0)$ (steps 2 and 3 in Algorithm I). The adversary will then break the links in $\mathcal{S}$ and will spare the link corresponding to $\mathcal{L}_{\pi,\ell}^\ell(0)$ as required. To see why this is optimal, consider the type of links that can be in $\mathcal{S}$.

- If a link in $\mathcal{S}$ is also in $\mathcal{L}_\pi^\ell(0)$, then this is optimal due to the fact that the adversary will disconnect that link since it is in $\mathcal{L}_\pi^\ell(0)$. Hence, if the designer can utilize this link to modify the ranking and protect a link that is more negative ($\mathcal{L}_{\pi,\ell}^\ell(0)$ in this case), then this can only improve his utility. The same reasoning applies if more than one of the links in $\mathcal{S}$ are also in $\mathcal{L}_\pi^\ell(0)$.
- If none of the links in $\mathcal{S}$ is in $\mathcal{L}_\pi^\ell(0)$, then necessarily some of the links in $\mathcal{L}_\pi^\ell(0)$ will also be protected along with the link corresponding to $\mathcal{L}_{\pi,\ell}^\ell(0)$. This is because $|\mathcal{S}| = \ell$, and the adversary can only break ℓ links. Hence, this scenario is more favorable to the designer than the previous one and can therefore only improve his utility.

If such an $\mathcal{S}$ exists, then the designer has exhausted all possible moves, since $|\mathcal{S}| = \ell$, and the algorithm terminates (step 4 of the algorithm). Otherwise, If no such set exists in $\mathcal{P}(0)$, then the designer should try to protect the next most negative link whose value is precisely $\mathcal{L}_{\pi,\ell-1}^\ell(0)$ by finding a set $\mathcal{S}$ of of size $\ell - 1$. Since $\mathcal{L}_{\pi,\ell-1}^\ell(0) \geq \mathcal{L}_{\pi,\ell}^\ell(0)$, the link corresponding to $\mathcal{L}_{\pi,\ell}^\ell(0)$ along with $\mathcal{S}$ will constitute the set of ℓ links that the adversary will break. Then, the designer should set $v_{ij} = b$ for all the links in $\mathcal{S}$, and for the remaining action the designer should select the link with the most negative f_{ij} that is *not* in $\mathcal{L}_\pi^\ell(v_\mathcal{S}(b))$; this is precisely the set $\Phi_1\left(\overline{\mathcal{P}}(v_\mathcal{S}(b))\right)$ (step 3 of the algorithm). The reason behind searching in $\overline{\mathcal{P}}(v_\mathcal{S}(b))$ and not in $\mathcal{P}(v_\mathcal{S}(b))$ after finding $\mathcal{S}$ is that the a_{ij}'s only affect the utility of the designer when he attempts to alter the ranking. Otherwise, the edge weights do not affect the utility of the minimizer as per (7).

This procedure then repeats until the designer has tried to protect all the links in $\mathcal{L}_\pi^\ell(0)$. If the designer fails in protecting *all* the links in $\mathcal{L}_\pi^\ell(0)$, then we must have $N_v = 0$, i.e., the input strategy was not altered. Then, the optimal strategy is to set $v_{ij} = b$ for the links with most negative f_{ij}'s in $\overline{\mathcal{P}}(0)$ (steps 7 and 8 in Algorithm I). ∎

C. Complexity of the Optimal Strategies

Let us now study the complexity of the optimal strategies. We first start by the max–min problem. Assuming as in Remark 1 that the players switch their strategies a total of K times over $[0, T]$, we conclude that the worst-case complexity of the strategy of both players is $\mathcal{O}(K \cdot m \log m)$ as their strategies involve merely the ranking of sets of size at most $2m$. As for the min–max problem, the complexity of the adversary's strategy is $\mathcal{O}(K \cdot m \log m)$. The main bottleneck in the strategy of the designer is step 2 in Algorithm I. The size of the set $\mathcal{P}(0)$ is at most $m - \ell$; thus, the worst-case complexity for the designer is $K \cdot \sum_{i=1}^{m-\ell} \binom{m-\ell}{i} \approx K \cdot \sum_{i=1}^{\ell} (m - \ell)^i$. By comparison with (2), we conclude that the derived optimal strategies achieve vast complexity reductions.

IV. IMPLEMENTING THE OPTIMAL STRATEGIES

The optimal strategies we derived in Section III are defined in terms f_{ij}'s which depend on the state $x(t)$ and the costate $p(t)$. However, we have not derived the optimal trajectories that satisfy the canonical equations given by the MP in (4) and (5). Since the system is linear time-varying, the solutions will be given in terms of a state transition matrix; this makes working with f_{ij} intractable. The following theorem provides a procedure to arrive at the optimal solutions without the need to compute $p(t)$. Parts of the proof of this theorem appeared in [12]. Define $\nu_{ij} = -(x_i - x_j)^2$.

Theorem 3: The rankings performed as part of the optimal strategies of the max–min and min–max problems can be carried out by replacing $f_{ij}(t)$ by $\nu_{ij}(t)$. Further, it is optimal for the players to modify a total of ℓ links.

Proof: We will provide the main ideas behind proving this claim for the max–min problem – the same arguments will apply to the min–max problem. For a fixed strategy of the adversary $u^\star$, we will show that it is optimal for the minimizer to rank the links based on their ν_{ij} values instead of the f_{ij}'s. Then, by applying the same reasoning as in Thm 1, we arrive at $u^\star$ with the ranking performed using ν_{ij}.

The main complication in solving the adjoint equations is that the system is time-varying. However, since $x(t), p(t) \in C^1[0, T]$, the value of f_{ij} cannot change abruptly in a finite interval. As a result, the control obtained from the MP cannot switch infinitely many times in a finite interval. Hence, we can regard the system as a time-invariant one over a small interval $[t_0, t_0 + \delta] \subset [0, T]$, $\delta > 0$. The proof consists of three steps.

1. Show that it is optimal for the players to change ℓ links.
2. Show that, over a small interval $[t_0, t_0 + \delta]$, it is optimal for the designer to switch from a strategy v^A to a strategy v^B that entails ranking the links based on their ν_{ij} values.

3. Show that mimicking v^A for the remaining time of the problem preserves the gain obtained over $[t_0, t_0 + \delta]$.

The first step was proven in [12] for the maximization problem. In the min–max (or max–min) case, the proof remains the same. This is because for a fixed $u^\star$, then by Thm 1 in [12], it follows that it is optimal for the designer to modify ℓ links[1]. By Thm 1, we deduce that $N_{u^\star} = \ell$.

We want to show that the designer will modify the ℓ links with the lowest ν_{ij} values. Assume that the designer applies a stationary strategy v^A over the interval $[t_0, t_0 + \delta]$ with a corresponding system matrix A. Let $P(t) := e^{At}$. Due to the structure of A, $P(t)$ is a doubly stochastic matrix for $t \geq 0$. Because the control strategies are time-invariant, the state trajectory is given by

$$x(t) = e^{A(t-t_0)}x(t_0), \quad t \in [t_0, t_0 + \delta].$$

Assume that the links the designer modifies over this interval are not the ones with the lowest ν_{ij} values. In particular, assume that the designer sets $v_{ij}^A = 0$ and $v_{kl}^B = b$ although $\nu_{ij} < \nu_{kl}$. At time $t^\star \in [t_0, t_0 + \delta]$, we assume that the designer switches to strategy v^B by setting $v_{ij}^B = b$ and $v_{kl}^B = 0$. Let the matrix B be the system matrix corresponding to v^B, and define the doubly stochastic matrix $Q(t) := e^{Bt}$, $t \geq 0$. Showing that this switch can improve the utility of the designer is equivalent to proving the following inequality [12]:

$$\int_{t^\star}^{t_0+\delta} k(t) \cdot x(t_0)^T \Lambda(t, t^\star) x(t_0) dt < 0, \quad (8)$$

where $\Lambda(t, t^\star) = P(t^\star - t_0)Q(2(t-t^\star))P(t^\star - t_0) - P(2(t - t_0))$. A sufficient condition for (8) to hold is

$$h(t, x(t_0)) = x(t_0)^T \Lambda(t, t^\star) x(t_0) < 0, \text{ for } t > t^\star.$$

As $t \downarrow 0$, we can write $P(t) = I + tA + \mathcal{O}(t^2)$, where $\mathcal{O}(t^2)/t \leq L$ for some finite constant L. Following similar steps to those in [12], we can write

$$h(t, x(t_0)) = 2(t - t^\star) \sum_{r>s}(A_{sr} - B_{sr})(x_r(t_0) - x_s(t_0))^2$$
$$= 2b(t - t^*)(\nu_{ij}(t_0) - \nu_{kl}(t_0)) + \mathcal{O}(\delta^2).$$

For small enough δ, the higher order terms are dominated by the first term. Hence, if there are links $\{i, j\}$ and $\{k, l\}$ such that $\nu_{ij} \leq \nu_{kl}$, there exists $t^\star$ such that $h(t, x(t_0)) < 0$ for $t \in (t^\star, t_0 + \delta]$. This proves that the optimal strategy for the designer is to break the links with the lowest ν_{ij} values. By comparison with the statement of Thm 1, we conclude that we can replace f_{ij} by ν_{ij} while performing the required rankings.

The final step of the proof is to show that switching to strategy B guarantees an improved utility for the adversary *regardless of how the original trajectory A changes beyond time $t_0 + \delta$*. To this end, we will assume that from time $t_0 + \delta$ onward, strategy B will mimic strategy A. Assume that strategy A switches to another strategy C starting at time

[1]Conditions were given in [12] as to when it is also optimal for a player to change less than ℓ links.
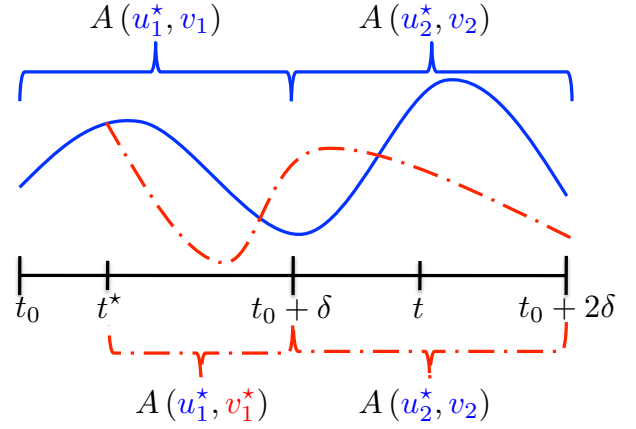

Fig. 2. A demonstration of the technique used in the third step of the proof.

$t_0 + \delta$. Hence, strategy B will also switch to strategy C, but the trajectories will have a different initial conditions. Fig. 2 illustrates this idea. Consider the behavior of the system over $[t_0 + \delta, t_0 + 2\delta]$ over which we can assume that the system is time-invariant. One can show that it suffices to prove that [17]

$$\int_{t_0+\delta}^{t_0+2\delta} k(t) \cdot c(t)dt < 0,$$

where $c(t)$ is the difference between the utilities resulting from the trajectory that includes a switch to strategy B at time $t^\star$ and the one that does not switch at time $t^\star$. After performing a first-order Taylor expansion, $c(t)$ can be shown to be equal to [17]

$$c(t) = 2b(t_0 + \delta - t^\star)(\nu_{ij}(t_0) - \nu_{kl}(t_0)) + \mathcal{O}(\delta^2),$$

which is negative for small δ. Hence, the gain that was obtained by switching to strategy B at time $t^\star$ is preserved and the proof is complete. ∎

Remark 3: *(Potential-Theoretic Analogy)* As discussed in [12], when the graph is viewed as an electrical network, $a_{ij} + v_{ij}$ can be viewed as the conductance of link $\{i, j\}$ and $x_i - x_j$ as the potential difference across the link. Therefore, the optimal strategy of the designer in both problems involves finding the links with the highest potential difference (or the lowest ν_{ij}'s) and increasing the conductance of those links by setting $v_{ij} = b$. This leads to increasing the power dissipation across those links, which translates to increasing the information flow across the network and results in faster convergence. The optimal strategy of the adversary should therefore involve breaking the links with highest power dissipation. But power dissipation is given by $(a_{ij} + v_{ij})(x_i - x_j)^2$, and this is exactly what the adversary targets according to Thm 3.

V. A SUFFICIENT CONDITION FOR THE EXISTENCE OF AN SPE

Having solved the min–max and max–min problems, it remains to verify whether the value of these problems can

be equal, even though the strategy sets of the players are non-rectangular. The reason that the upper and lower values are different is mainly due to the ability of the minimizer to *deceive* the maximizer by altering the ranking of the most negative values. If we remove this ability from the network designer, we should expect that an SPE would exist. The following theorem makes this argument formal. Define $\gamma := \frac{M}{\epsilon^2}$, where M, ϵ are defined in Lemmas 1 and 2 in the Appendix. We assume that ϵ is chosen to guarantee $\gamma > 1$.

Theorem 4: A sufficient condition for the upper and lower values $\overline{V}, \underline{V}$, defined in (3) to be equal, and hence for the existence of an SPE, is to select b such that

$$0 \le b \le \min_{\{i,j\},\{k,l\}\in\mathcal{E}} |\gamma a_{ij} - a_{kl}|, \tag{9}$$

given that $a_{ij} \ne a_{kl}$ and $a_{ij} > \gamma a_{kl}$ whenever $a_{ij} > a_{kl}$, for all $\{i,j\}, \{k,l\} \in \mathcal{E}$.

Proof: It suffices to show that $\mathcal{L}_\pi^\ell(v^\star) = \mathcal{L}_\pi^\ell(0) = \mathcal{D}_\pi^\ell(G^\star)$ as this would imply that the adversary would break the same links whether he acts first or second, and as a result the strategy of the minimizer in both problems will be the same. This would occur if the minimizer cannot protect any of the links in $\mathcal{L}_\pi^\ell(0)$. In other words, this will occur if the minimizer cannot satisfy the condition in step 2 of Algorithm I for any $1 \le i \le \ell$. A sufficient condition for this to happen is to require

$$\min_{\{i,j\}\in\mathcal{P}(0)} (a_{ij} + b) f_{ij} > \max_{\{i,j\}\in\mathcal{L}_\pi^\ell(0)} a_{ij} f_{ij}.$$

This implies that no matter how the designer changes the weights of the links in $\mathcal{P}(0)$, it cannot make those links more negative than the links in $\mathcal{L}_\pi^\ell(0)$. To satisfy this inequality, we will establish that whenever $a_{ij} f_{ij} > a_{kl} f_{kl}$, $f_{ij}, f_{kl} < 0$, we must have $(a_{ij} + b) f_{ij} > a_{kl} f_{kl}$, for all links in $\mathcal{E}$. We can then re-write the condition on b as

$$b \le \frac{a_{ij} f_{ij} - a_{kl} f_{kl}}{-f_{ij}} = a_{kl} \frac{|f_{kl}|}{|f_{ij}|} - a_{ij}. \tag{10}$$

Consider the following two cases. If $f_{kl} \ge f_{ij}$, then we must have $a_{kl} > a_{ij}$. Then, by assumption we have that $a_{kl} > \gamma a_{ij}$. By Lemmas 1 and 2, we can write

$$a_{kl} \frac{|f_{kl}|}{|f_{ij}|} - a_{ij} \ge \frac{1}{\gamma} a_{kl} - a_{ij} > 0. \tag{11}$$

Next, consider the case when $f_{ij} > f_{kl}$. In this case, a_{ij} can be greater or smaller than a_{kl}. However, if $a_{ij} > a_{kl}$, and recalling that $a_{ij} f_{ij} > a_{kl} f_{kl}$, then

$$\gamma a_{kl} < a_{ij} < a_{kl} \frac{|f_{kl}|}{|f_{ij}|} \le \gamma a_{kl},$$

which is a contradiction. The case $a_{kl} = a_{ij}$ is excluded by assumption. Hence, in this case, we must have $a_{ij} < a_{kl}$, and the inequality in (11) applies. Thus, by choosing b as in (9), we obtain the condition we are seeking. Note that we do not need to consider the case when $a_{ij} f_{ij} = a_{kl} f_{kl}$ since the players will be indifferent as to which link to choose when the corresponding values are equal. ∎

Remark 4: The condition derived in the above theorem requires the network to be *sufficiently diverse* in the sense that the weights of the links have to be not only different from each other, but also a factor γ apart. This is due to the fact that we were seeking uniform bounds on the f_{ij}'s. If we allow b to vary with time, then one can find less restrictive conditions to ensure the existence of an SPE. However, this would require (10) to be verified at each time instant. Further, the bounds derived in the Appendix are loose. Tighter bounds could be given by studying the dynamics of $|f_{kl}|/|f_{ij}|$. However, studying the time derivative of this ratio is not tractable.

VI. Conclusion

In this paper, we have formulated and solved two problems that capture the competition between an adversary and a network designer in a dynamical network. We considered practical models for the players by constraining their actions along the problem horizon. The derived strategies were shown to exhibit a low worst-case complexity. We also proved that the optimal strategies admit a potential-theoretic analogy, and they can be implemented without needing to solve the adjoint equations. Finally, we showed that when the link weights are sufficiently diverse an SPE exists for the zero-sum game between the designer and the adversary.

References

[1] R. Olfati-Saber and R. M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Trans. Automat. Contr.*, vol. 49, no. 9, pp. 1520–1533, 2004.

[2] R. Olfati-Saber, "Flocking for multi-agent dynamic systems: Algorithms and theory," *IEEE Trans. Automat. Contr.*, vol. 51, no. 3, pp. 40–420, 2006.

[3] A. Nedić, A. Ozdaglar, and A. Parrilo, "Constrained consensus and optimization in multi-agent networks," *IEEE Trans. Automat. Contr.*, vol. 55, no. 4, pp. 922–938, 2010.

[4] M. O. Jackson and B. Golub, "Naive learning in social networks: Convergence, influence and wisdom of crowds," *American Economic J.: Microeconomics*, vol. 2, no. 1, pp. 112–149, 2010.

[5] A. Nedić and A. Ozdaglar, "Convergence rate for consensus with delays," *J. Global Optimization*, vol. 47, no. 3, pp. 437–456, 2010.

[6] L. Xiao, S. Boyd, and S.-J. Kim, "Distributed average consensus with least-mean-square deviation," *J. Parallel and Distributed Computing*, vol. 67, no. 1, pp. 33–46, 2007.

[7] B. Touri and A. Nedić, "Distributed consensus over network with noisy links," in *Proc. 12th Internat. Information Fusion Conf.*, 2009, pp. 146–154.

[8] A. Kashyap, T. Başar, and R. Srikant, "Quantized consensus," *Automatica*, vol. 43, no. 7, pp. 1192–1203, 2007.

[9] S. Sundaram and C. Hadjicostis, "Distributed function calculation via linear iterative strategies in the presence of malicious agents," *IEEE Trans. Automat. Contr.*, vol. 56, no. 7, pp. 1495–1508, July 2011.

[10] A. Teixeira, H. Sandberg, and K. Johansson, "Networked control systems under cyber attacks with applications to power networks," in *Proc. American Control Conference*, July 2010, pp. 3690–3696.

[11] M. Pease, R. Shostak, and L. Lamport, "Reaching agreement in the presence of faults," *Journal of the ACM*, vol. 27, no. 2, pp. 228–234, 1980.

[12] A. Khanafer, B. Touri, and T. Başar, "Consensus in the presence of an adversary," in *3rd IFAC Workshop on Distributed Estimation and Control in Networked Systems (NecSys)*, September 2012, pp. 276–281.

[13] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge university press, 2005.

[14] S. Goyal and A. Vigier, "Robust networks," mimeo, University of Cambridge, Faculty of Economics, Tech. Rep., 2010.

[15] H. J. LeBlanc, H. Zhang, S. Sundaram, and X. Koutsoukos, "Consensus of multi-agent networks in the presence of adversaries using only local information," in *Proc. 1st Internat. Conf. High Conf. Net. Sys.*, 2012, pp. 1–10.

[16] F. Pasqualetti, A. Bicchi, and F. Bullo, "Consensus computation in unreliable networks: A system theoretic approach," *IEEE Trans. Automat. Contr.*, vol. 57, no. 1, pp. 90–104, January 2012.

[17] A. Khanafer, B. Touri, and T. Başar, "Robust distributed averaging in networks," *http://arxiv.org/abs/1303.2422*, March 2013.

APPENDIX

Lemma 1: The function $f_{ij}(t)$ is uniformly bounded from above by

$$M := 4nT \left(||x(0)||_\infty + |x_{avg}| \right) ||x(0)||_\infty .$$

Proof: By the structure of the system matrix in (1), we can deduce that $|x_i - x_j|$ cannot increase as $t \to T$. Thus

$$
\begin{aligned}
|x_i(t) - x_j(t)| &\leq \max_{1 \leq i,j \leq n} |x_i(0) - x_j(0)| \\
&\leq 2 \max_{1 \leq i \leq n} |x_i(0)| = 2 ||x(0)||_\infty .
\end{aligned}
$$

In order to bound $|p_i(t) - p_j(t)|$, we consider the interval $[T - \delta, T]$ where $\delta > 0$ is small enough such that the system in (1) is time-invariant. Let the system matrix over this interval be A and define $s := T - \delta$. The solution to the adjoint dynamics in (4), (5) over $[s, T]$ is then given by

$$x(t) = e^{A(t-s)} x(s), \tag{12}$$

$$p(t) = 2 \int_t^T e^{-A(t-\tau)} (x(\tau) - \bar{x}) d\tau. \tag{13}$$

Define the doubly stochastic matrix $P(t) := e^{At}$. Then

$$
\begin{aligned}
|p_i - p_j| &= 2 \left| \int_t^T (P_i(\tau - t) - P_j(\tau - t))(x(\tau) - \bar{x}) d\tau \right| \\
&\leq 2 \int_t^T \sum_{k=1}^n |(P_{ik}(\tau - t) - P_{jk}(\tau - t))(x_k(\tau) - \bar{x}_k)| \, d\tau \\
&\leq 2 \int_t^T \sum_{k=1}^n |x_k(\tau) - \bar{x}_k| \, d\tau \\
&\leq 2 \int_t^T (||x(\tau)||_1 + n|x_{avg}|) d\tau.
\end{aligned}
$$

Note that over $[s, T]$ we have $||x(t)||_1 = ||P(t-s)x(s)||_1 \leq ||P(t-s)||_1 ||x(s)||_1 = ||x(s)||_1$, where the last inequality follows from the fact that $P(t)$ is a stochastic matrix. By repeating the above argument over the interval $[0, s)$, after dividing it into smaller intervals and exploiting the fact that the solution over each small interval will be of the same form as in (12), we conclude that $||x(t)||_1 \leq ||x(0)||_1$, for all $t \in [0, T]$. We can therefore write

$$|p_i(t) - p_j(t)| \leq 2\delta (||x(0)||_1 + n|x_{avg}|), \quad t \in [s, T].$$

When we propagate the solution backward over the interval $[T - 2\delta, \delta]$, a new integral term will be added to the expression in (13); however, it will be of the same structure as that of $p(t)$ over $[s, T]$: an integral of a doubly stochastic matrix that depends on the system matrix over $[T - 2\delta, T - \delta]$ multiplied by $x(t) - \bar{x}$. This applies for all infinitesimal intervals we consider from $T - 2\delta$ down to zero. Hence, if we subdivide $[0, T]$ into K intervals of length δ and apply the triangular inequality to $|p_i(t) - p_j(t)|$, we will obtain

$$|p_i(t) - p_j(t)| \leq 2T (||x(0)||_1 + n|x_{avg}|), \quad t \in [0, T].$$

Finally, recall that $||x(0)||_1 \leq n ||x(0)||_\infty$. We therefore have

$$|f_{ij}| = |p_i - p_j||x_i - x_j| \leq 4nT (||x(0)||_\infty + |x_{avg}|) ||x(0)||_\infty ,$$

as claimed. ∎

Lemma 2: Given δ, $\epsilon > 0$, one can select the problem horizon T small enough such that $|f_{ij}| \geq \epsilon^2$ for all $t \in [0, T]$.

Proof: We now show that $|f_{ij}(t)|$ can be bounded below uniformly by a positive number. In order to do so, we need to ensure that $|x_i(t) - x_j(t)|$ does not approach zero as $t \to T$. We are seeking a time $t^\star$ such that for a given $\epsilon > 0$, we have $|x_i(t) - x_j(t)| \geq \epsilon$ for all $t < t^\star$ and all i, j. We can then fix $T < t^\star$ to ensure the existence of a uniform lower bound on $|x_i(t) - x_j(t)|$. Let us again restrict our attention to a small interval $[t_0, t_0 + \delta]$ where the system is time-invariant, and let the system matrix over this interval be A. We require that the system did not reach equilibrium over this interval, i.e., $x(t_0 + \delta) \neq \bar{x}$. Without loss of generality, we assume that $x_1(t_0) > \ldots > x_n(t_0)^2$. Define the following dynamics

$$
\begin{aligned}
\frac{d}{dt}(\bar{y}_i - x_1(t_0)) &= \sum_{j \neq i} A_{ij}(x_1(t_0) - \bar{y}_i), \\
\frac{d}{dt}(\underline{y}_i - x_n(t_0)) &= \sum_{j \neq i} A_{ij}(x_n(t_0) - \underline{y}_i),
\end{aligned}
$$

with initial conditions $\bar{y}_i(t_0) = 2x_1(t_0)$, $\underline{y}_i(t_0) = 2x_n(t_0)$. Note that $\dot{x}_i = \sum_{j \neq i} A_{ij}(x_j - x_i)$. It follows that $\dot{\underline{y}}_i \leq \dot{x}_i \leq \dot{\bar{y}}_i$. By the comparison principle, we conclude that $\underline{y}_i - x_n(t_0) \leq x_i \leq \bar{y}_i - x_1(t_0)$, for $1 \leq i \leq n$. Note that we can readily find the solution trajectories for $\bar{y}$ and $\underline{y}$. By defining $a_i = \sum_{j \neq i} A_{ij}$, we can then write

$$
\begin{aligned}
\bar{y}_i - x_1(t_0) &= e^{-a_i(t-t_0)} x_1(t_0), \\
\underline{y}_i - x_n(t_0) &= e^{-a_i(t-t_0)} x_n(t_0).
\end{aligned}
$$

By solving the equation $\bar{y}_{i-i} - x_1(t_0) = \underline{y}_i - x_n(t_0)$, we can find a time $t_i^\star$ when x_{i-1} can potentially meet x_i:

$$t^\star = \frac{1}{a_{i-1} - a_i} \ln \left(\frac{x_1(t_0)}{x_n(t_0)} \right) + t_0.$$

If $t_i^\star > t_0 + \delta$, for all i, then we need to propagate the solution forward, and keeping in mind that the system matrix could change, until we find a time $t_i^\star$ in some interval $[\tilde{t}, \tilde{t} + \delta]$ where $\bar{y}_{i-i} = \underline{y}_i$ for some i. Then, for a given $\epsilon > 0$, we can select $T < t^\star$ such that $|x_i - x_{i-1}| \geq |\underline{y}_i - \bar{y}_{i-1}| \geq \epsilon$; hence, we conclude that for this choice of T we can guarantee that $|x_i - x_j| \geq \epsilon > 0$ for all i, j.

Since $p(T) = 0$, then $|p_i - p_j| \to 0$ as $t \to T$. Given $\epsilon > 0$, there is a time $\tilde{T}_{ij}(\epsilon) < T$ such that $|p_i - p_j| < \epsilon$, for all $t \geq \tilde{T}_{ij}(\epsilon)$. Define the time $\tilde{T} = \min_{1 \leq i,j \leq n} \tilde{T}_{ij}(\epsilon)$. Then $|p_i - p_j| \geq \epsilon$ for $t \geq \tilde{T}(\epsilon)$, for all i, j. We therefore have $|f_{ij}| = |p_i - p_j||x_i - x_j| \geq \epsilon^2 > 0$. ∎

---

[2] We are making the implicit assumption that $x_1(0) > \ldots > x_n(0)$.